Journal of Biomedical Informatics **34,** 321–347 (2001)
doi:10.1006/jbin.2001.1030

Symbolic Representation of Anatomical Knowledge: Concept Classification and Development Strategies

P. Cerveri*,†,[1] and F. Pinciroli*

**Bioengineering Department, Politecnico di Milano, Piazza Leonardo da Vinci, 32, I-20133 Milan, Italy; and*
†Centro di Bioingegneria, Politecnico di Milano–Fondazione Pro Juventute Don Carlo Gnocchi ONLUS, Via Capecelatro, 66, 20148 Milan, Italy

Received July 25, 2001; published online May 23, 2002

In this paper a novel approach to anatomy knowledge representation is described. The focus of the present research has been on the development of a representational framework where the conceptual level has been implemented by using hierarchical and nonhierarchical conceptual networks. This has allowed handling the demand for multiple views of anatomy (systemic and topographical views). The terminological level of the knowledge representation has been implemented by using a compositional strategy which has avoided the explicit storage of the terms used to express composite concepts. Hierarchical relations and composite concept representations have required supervision of both the inheritance and concept reconstruction. For this purpose heuristic knowledge has been stored in terms of consistency rules in the knowledge base. As proof of the capability of this system, we show how the knowledge base has been used to provide symbolic access to spatial information consisting of a reduced set of images from the Visible Human Dataset.

1. INTRODUCTION

Despite a long history of research in the field of medical informatics, the design and implementation of accurate, complete, and flexible representations of biomedical knowledge remain an as yet unfulfilled goal. Among the causes of this deficiency, one reason appears to be the minimal attention paid, by both knowledge engineers and software developers, to the problem of adequately modeling human anatomy knowledge in computer-based systems. In fact, a large number of the assertions formulated in all biomedical domains make use of anatomical concepts. Clinical treatments, diseases, biochemical processes, and surgical interventions all imply the generation of statements in which the involved concepts refer to body locations, organs, or generic body components at macroscopic as well as microscopic levels. Endowing a computer system with the ability to understand biomedical statements and perform reasoning about that content—what we intend to clarify here with reasoning—is not a trivial task. Let us consider the following phrases as typical examples of clinical statements involving anatomical concepts:

- Cirrhosis affects the liver or more precisely a part of the liver, manifests in its interior, and cannot affect body parts other than liver,
- Disease affects an organ or organ part, manifests in the interior or exterior part of the organ, and always affects the same organ type.

The first statement refers to a specific disease affecting

[1]To whom correspondence and reprint requests should be addressed. E-mail: cerveri@biomed.polimi.it.

1532-0464/01 $35.00

a particular organ (real concepts). In contrast, the second deals with different levels of knowledge involving conceptual categories (abstract concepts). While the first statement implies specific knowledge of anatomical entities (special anatomy), the second requires some *a priori* assumptions used to classify concepts into categories (general anatomy). With the aim of emulating the human mind's capability to understand both levels of knowledge, a computer-based system should be provided with four main frameworks: (1) a concept classification, (2) data structures (knowledge base) into which concepts are mapped, (3) a terminological source to map concepts to language, and (4) a software engine to perform reasoning (knowledge-based querying and information reconstruction).

Terminological sources, widely used in biomedicine, represent a low-level attempt to organize biomedical knowledge into a computer-based system. The aim of such sources is twofold: (a) they endeavor to achieve standardization of the terms used in a specific domain of knowledge; (b) they attempt to provide a symbolic representation of underlying concepts [2–4]. A *terminology* is a low-level system that aggregates terms according to simple alphabetical rules; however, no assumptions about the conceptual organization are made. By assigning unique reference codes to terms, a terminology can be made into a *coding system*. Furthermore, it can be made into a *thesaurus* by distinguishing between preferred terms and synonyms. Any preferred term constitutes the main reference for a concept, whereas the corresponding synonyms are sometimes used to designate the same concept. By providing definitions for the terms, a thesaurus assumes the characteristics of a *vocabulary*. However, a vocabulary is a static system and cannot be queried to validate any semantic statement.

With regard to anatomy, several terminological sources which collect terms used by experts in specific clinical environments exist. Depending on the medical field in which they have been used, the same terms can sometimes refer to different anatomical concepts—leading to inconsistency between different sources, or multiple different terms can refer to a single concept–causing redundancy. The Unified Medical Language System (UMLS) [5] from the National Library of Medicine was conceived as an attempt to build a standardization interface among terminological sources, with the aim of removing term redundancy. As a result, the Metathesaurus, the main knowledge source of UMLS, makes available alphanumeric codes that act as links between concepts and terms in individual terminological sources. The UMLS does not provide a proper terminological source itself, but rather adopts those of the terminology providers. Therefore, while eliminating redundancy, it does not resolve term inconsistency. Only moving from a terminological to a conceptual level can address the problem of term inconsistency.

Terminological sources as referred to here are considered to be simple coding systems containing lists of coded terms but without any explicit relations among the corresponding concepts. Actually, the available terminological sources provide term organization but little concept classification. Semantic networks or conceptual graphs are the primary visualization tools used to represent concepts and their interrelations [6, 7]. In semantic networks, knowledge is represented by nodes (concepts) and arcs (semantic relations). In particular, semantic relations divide into hierarchical and nonhierarchical relations. Starting with most generic concepts, concept trees can be obtained in which specialization increases as the leaves of the tree are approached. For example, relating concepts through relations of genus (::*is-a-kind-of*) allows the generation of a taxonomic tree. Equivalently, a partonomy tree can be obtained by linking concepts through a part-of relation. Nonhierarchical relations can be used to model other concept attributes.

With respect to concept classification, the UMLS provides a knowledge source named the Semantic Network. However, its model takes into account only a few semantic types that do not allow for anatomical entities to be adequately represented. In SNOMED [8], only strict concept classification is provided. A simple nonpersistent strict hierarchy based on topographical properties of anatomical objects is realized by using significant alphanumeric codes assigned to terms. Although it collects a relatively large number of anatomical terms, this source exhibits insufficient flexibility for accurate anatomy knowledge representation. The Read Codes Project [9] pursues a more flexible approach in which concepts, separated and linked to terms by nonsignificant codes, have been classified into categories and structured into a taxonomic semantic network. The system, focusing on anatomical structures, has been conceived to express several taxonomic views but does not provide either partonomic views or establish nonhierarchical relations between concepts. For example, the cribriform plate, being a part of the lateral mass (left and right) of the ethmoid bone, is coded generically as a bone structure of the cranium without any further specification. However, while expressing adequate representational power, albeit with some inconsistencies in the anatomical representation, the project has the great advantage of addressing the problem of compositional organization of concepts. In antithesis to enumeration, composition implies that terms indicating composite concepts are implicitly stored in the knowledge base.

The GALEN project [10–12] adopts, as other systems do, enumeration for representing anatomical entities, but also provides a compositional framework (the CORE model) manageable by a flexible language named GRAIL [13]. These two modules guarantee the ability to derive, from a reduced set of concepts, articulated concepts and phrases that are consistent and nonredundant due to automated rules. However, the anatomical terminology expressed in GALEN has the drawback of assuming anatomical entities as sites of disease processes, which reduces the ability to represent fully detailed levels of anatomical objects.

Unlike the project of Rosse's group [14], none of the above initiatives has systematically addressed the problem of building anatomical concept classification in the realm of pure anatomy. The aim of that project has consisted of defining a foundational model for anatomy able to connect a concept classification to a source of terms and to accommodate a large typology of anatomical views. Much effort has been spent to develop an accurate anatomical terminology (about 25,000 terms), and concept definitions have been generated consistently with concept classification properties. The approach pursued to collect concepts has been the enumerative strategy, and both multiple hierarchies and nonhierarchical relations between concepts have been underestimated.

The approach of the Höhne group [15] was focused on the representation of head anatomical structures and their interrelations. The great advantage of the methodology applied is that it addresses the problem of the partonomic and nonhierarchical relations. This methodology has involved the use of nonhierarchical relations to describe characteristics of nerves and vessels. However, multiple concept classifications and inheritance have been disregarded.

We note here that no report in the cited literature has given consideration to inheritance as a basis for the classification. In principle, the property of inheritance of hierarchical relations guarantees that attributes can be automatically passed down from parent to child (concept of monotonic inheritance). This allows distribution of semantic attributes of the concepts along one or more hierarchical networks constituted by conceptual categories. In practice, feature inheritance is very composite and full of exceptions and peculiarities (nonmonotonic inheritance). This implies that at each level of a hierarchy the attempt to subsume attributes must be accurately validated by consistency rules. This issue has received little attention to date.

In light of these considerations, our work has been focused on developing an anatomical knowledge model, based on concept hierarchies, endowed with an information recovery engine to constrain inheritance. In particular, it differs from previous approaches with respect to the following features:

Both hierarchical and nonhierarchical relations between concepts have been used.

Polyhierarchies have been introduced to allow maximum expressiveness as required by anatomical knowledge.

A reconstruction information engine has been used to obtain explicit knowledge from its implicit storage.

A supervised inference algorithm has been developed to reconstruct knowledge consistently blocking undue inheritance.

Conscious of the fact that we cannot coordinate all anatomical information content, we focused our conceptual framework on an orientation to beginners and assumed that our system can be progressively improved.

Moreover, among our aims, we did not set the goal of refining any specific terminology. Rather we attempted to eliminate term redundancy by allowing a real-time connection to the UMLS server able to map our terminology to other terminological sources. In the following sections we explain these characteristics.

2. MODEL AND METHODS FOR ANATOMICAL REPRESENTATION

2.1. Basis of Knowledge Representation

The aim of any representation is to organize the entities of a domain of knowledge according to some principles or commitments that specify how to look at the attributes of domain entities. In particular, these principles force one to distinguish what the significant information content entities exhibited and what must be accurately modeled from what is less relevant, and thus what can be disregarded or represented with less precision. This formal operation of conceptualization, supervised by some a priori assumptions (commitments), is named concept classification. In other words, concept classification foresees a subdivision into general and specific categories based on similarity and discriminating properties (intrinsic attributes) of domain objects (individuals/instances). Thus, representing knowledge consists of structuring concepts by expressing semantic constraints, namely semantic relations that have been identified.

In general, a relation is a function of one or more arguments that can be specified in a multivalued variable. For example, the concept muscle may be linked to the concept arm, with possible values that can include the flexion, adduction, torsion, etc., attributes. Equivalently, the concept skull may be linked to the concept bone through the composed-by relation. Binary or two-entity relations are specific relations that join two concepts in a semantic statement. A simple binary relation can be expressed as A::*r* B, which is asserting that entity A is related to B by relation *r*. Under this assumption, most of the properties of a concept A can be represented by a list of statements {A::*r*1 B, A::*r*2 C, A::*r*3 D, etc.} and graphically represented by a semantic network. The more that relations between concepts are identified, the more sophisticated the representation becomes. In particular, the use of binary relations allows creation of interconnected networks of concepts that are easily deliverable in a computer data structure. In Appendix A we report the definitions of the main semantic relationships referred to in the paper. However, they have standard definitions that can be retrieved by the UMLS knowledge base server.

The simplest anatomical conceptualization through a binary relation can be expressed by a statement such as hand::*is-a-kind-of* body part. Automatically the definition of the concept hand can be built as ‘hand is a body part.’ The problem here is what we mean by the concept body part. As reported by the UMLS, the following definition seems to be acceptable for the meaning of body part: “A collection of cells and tissues which are localized to a specific area or combine and carry out one or more specialized functions of an organism. This ranges from gross structures to small components of complex organs. These structures are relatively localized in comparison to tissues.” According to this definition, anatomical physical objects ranging from gross structures to organs and even small components of organs can be grouped together. As a consequence, any part of the body could be qualified as a body part, which does not appear to be useful. We are not in agreement with this definition because we think that the category body part should be assigned to only physical anatomical objects that can be externally discernible on the body, having virtual (external and internal) boundaries. The head can surely be an instance of the body part category: it is externally and internally separated from the thorax by the neck. In contrast, the concept right ventricle should not be subsumed by the body part category because an external subdivision of the body that contains it does not exist. However, in abstract terms it is a part of the body in the sense that the body includes it. In light of these considerations we propose two distinct definitions for the concept body part: one refers to the abstract concept as an aggregate of properties of anatomical entities that may pertain to content, surface, localization, or function, with at least one of them present; the other refers to the concrete concept of an anatomical structure as an aggregate of heterogeneous physical structures (organs, organ parts, and tissues) externally demarcated by skin subdivision. The latter has been used as a category in our representation and includes instances such as head, neck, shoulder, arm, and forearm.

2.2. Hierarchical Relations and Granularity Level

As stated, a classification consists of concept networks where relations link categories to categories and categories to individuals. In the case of monotonic inheritance, the transitive property of hierarchical relations (genus and partition) allows the distribution of concept properties across several hierarchical levels. In particular, each branch in a tree of categories represents a specification level at which a concept can be expressed. This aspect is strictly connected to the topic of granularity level of the description. Let us explain in more detail.

Medical and in particular anatomical information can be described and used according to a particular context of discourse. Detailed communications need to deal with a comprehensive representation of concepts; generic discourses are sufficiently accommodated by the coarse granularity level of the concepts used. For example, while discussing a serious injury, two clinical specialists might use a phrase like “a severe bilateral open vertical fracture of the sacrum with a complete cauda equina lesion.” In contrast, while communicating the patient’s condition to relatives, they would probably explain it simply as a bone fracture in the pelvis. The use of fine-grained concepts allows precise identification of objects but requires extensive knowledge to be understood, whereas coarse-grained ones, while losing details, guarantee immediate and intuitive insight. With this perspective, our representation has been developed with the aim of manipulating concepts at different levels of abstraction. For example, (see Fig. 1) the left and right ventricles can be categorized as the first approximation through the parent category ventricle. Moving to higher levels of abstraction, the ventricle can be classified as heart cavity and organ cavity.

In addition to the abstraction level, the detail level characterizes granularity of a concept. In particular for an anatomical concept, this refers to the description of entity subparts. The concept heart can be described as an anatomical entity with two main subdivisions, a left side and a right side. If more detail is needed, its definition can be refined to encompass information stating that the left side of the heart comprises the left ventricle and the left atrium, which are connected through the mitral valve. Furthermore, this definition can be improved. Both concept description and classification are based on some choice of granularity. The wider the granularity levels, the more clear and useful the representation, and the more it can capture the complexity of the knowledge domain. However, the wider the granularity levels, the more fragmented the knowledge and the more difficult it is to manage. To conclude, we note that a direct relation exists both between abstraction level and taxonomic classification and between detail level and partonomic classification. Moreover, hierarchical networks should be flexible in the sense of allowing one to collapse several abstraction levels, disregarding intermediate concepts, so that, for example, an individual can be viewed as directly linked to a high level category.

2.2.1. Is-a-kind-of relation. Let us now focus on the problem of the definition of categories by considering what standard anatomical books provide about the topic [17]. Some books present anatomy by defining the systems that constitute the human body (cardiovascular system, nervous system and so on) in such a way that the reader accesses knowledge through the systemic view. This represents a commitment assumed by book authors. In contrast, other books [17] provide a topographical view of anatomical content by describing the human body as constituted by regions and parts. As a result, merged views, which are not explicitly taken into account, involve effort by the reader to understand. Another notable aspect is the definition and description of concepts that books present. When describing a concept like the heart, one can ask what this concept is intrinsically. In general, each knowledge source provides an ad hoc answer to such questions. Organ, body part, muscle, involuntary muscle, and main structure of the cardiovascular system seem to be possible classes for heart. Yet, their meanings and their relationships may vary from one source to another. To overcome such inconsistency, we first focused our development on building high-level categories. It is well known that building a taxonomic tree requires linking of concepts by ::*is-a-kind-of* relations. All items in macroscopic anatomy can be easily assigned to three main categories: space, physical structure, and substance. From this reduced set of generic categories, anatomical concepts have been vertically specified through several abstraction levels (subsumption operation). The generation of horizontal categories, at a predetermined abstraction level, has been driven, where possible, by the principle of null intersection (complementary) classes.

Given the above classification, for example, the heart could clearly be associated with the physical structure category because of its intrinsic characteristics. However, to classify the heart, common sense suggests that organ is a category more specialized than physical structure. The definition of the fundamental category organ has been obtained according to the following statement: it is an anatomical structure; it is distinct both morphologically and functionally from other such units; it cannot be divided into organs; it is composed of tissue; it expresses several functions; it can be dense; it can contain hollow space; it can contain and produces substances; it can perform actions. Under these considerations, the following definition seems adequate: organ is a minimal self-contained anatomical structure able to express functional attributes. It is constituted by tissue that specifies its shape and its structural properties.

The following list is the formal representation of such a definition:

organ::*is-a-kind-of* anatomical structure
organ::*expresses* function
organ::*performs* action
organ::*is-constituted-by* tissue
organ part::*is-part-of* organ
full organ::*is-a-kind-of* organ
container organ::*is-a-kind-of* organ
organ::*is-part-of* body system

Based on the structural differences that organs exhibit, two major categories, container organ and full organ, have been generated. Container organs typically can enclose liquid, gas, or solid substance and necessarily are constituted by cavities. In contrast, full organs accomplishing the function of body sustainer (as bones), body motion (muscle), substance producer (liver, pancreas) are uniformly filled (approximately) by tissue. As an example of taxonomy, Fig. 2 shows the semantic network for the category container organ.

Note that we have identified three main subcategories of the category container organ: gas container, liquid container, and solid container organ. For container organ we have identified as relevant the morphological property of being of tubular shape which has led to the generation of the category tubular organ. This has led us to allow for the category tubular organ to have more than one parent concept. This is called a multiple-child/multiple-parent hierarchy (polyhierarchy), in contrast to strict hierarchies in which a child concept can have only one parent. As will become clearer in the following sections, polyhierarchies make the use of automated inheritance more problematic. Furthermore, it is important to emphasize here that the classification shown in Fig. 2 does not exhaust all structural and functional properties of container organs. For example, the concept lung has structural and functional properties that are not taken into account.

Based on specific macroscopic structural and functional

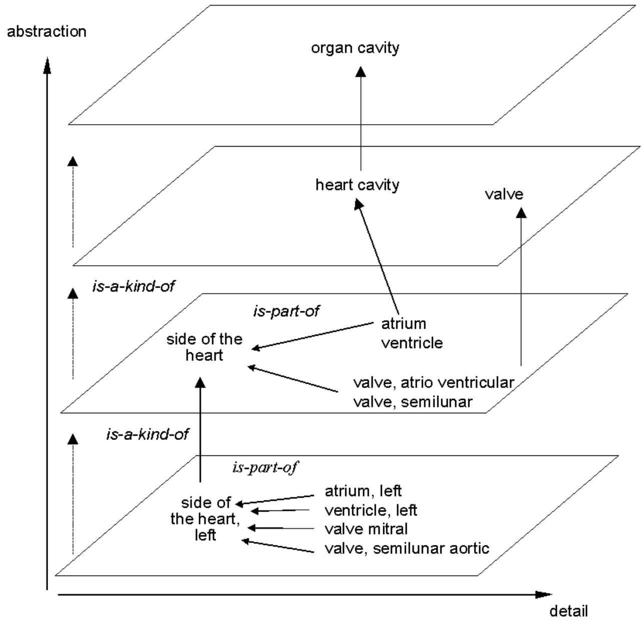

FIG. 1. Granularity: Abstraction and detail.

properties of full organs, we have designed categories like parenchymatous organ, glandular organ, bones, muscle, and joint, and for each of them subcategories have been generated.

For example, the concept bone has been subdivided into four main classes according to morphological properties (i.e., irregular bone, long bone, flat bone, and short bone). Each of these classes has been specialized according to topographical (i.e., cranial bone, carpal bone, vertebra) membership criteria. Finally, each single bone has been assigned to the category that defines it more specifically (see Fig. 3).

Similar criteria have been adopted to classify muscles and joints. In this case a strict hierarchy is not sufficient to express all properties that anatomical objects exhibit. For example, rib should be classified for its morphology as being a long bone. However, it shares an internal structure similar to flat bone (lack of marrow). In this case rib cannot have two parents (long and flat bone) because in the definition of category flat bone we did not take into account a structural criterion but only a morphological one. This property of the rib will be expressed differently as discussed here.

The above problem can be better illustrated in the effort to represent in the network the properties of the pancreas. It should be classified both as a parenchymatous organ like the liver and as a glandular organ like the thyroid gland. In addition, due to its specific functional properties, the pancreas could be classified as an endocrine gland (see Fig. 4). A similar issue relates to the classification of the lung as a container organ but also assigning the property that it is constituted by parenchymal tissue. In this case, however, the inheritance between lung concept and full organ concept is to be explicitly blocked.

An alternative solution consists of reducing a polyhierarchy into two or more strict hierarchies where inheritance is not a requirement. This approach can be illustrated by classifying muscles. At the first level, they can be classified according to fiber property, smooth and striated. For the second level, striated muscles have been specialized into cardiac muscle and skeletal muscle. Also it is reasonable to classify muscles into voluntary and involuntary. However, note that cardiac muscle and skeletal muscle are to be classified respectively as involuntary and voluntary, while sharing striated fibers. The attempt to take into account the first representation (smooth and striated muscle) for voluntary and involuntary muscle categories by associating them with both smooth and striated muscle would lead to the use of a constrained inheritance. This would produce a condition where cardiac muscle inherits both smooth and striated muscle properties via involuntary category. Therefore, this example shows that the genus relation can be characterized by a set of distinct contexts specifying multiple taxonomical views.

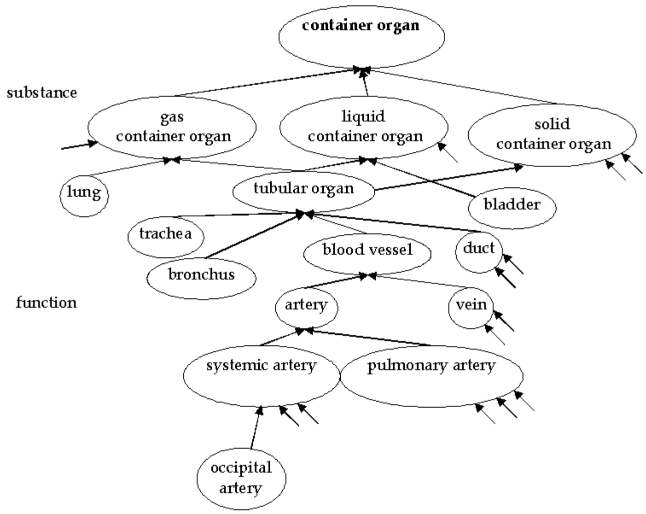

FIG. 2. The semantic network for category container organ. Any container organs can have a tubular shape. Note that this classification does not exhaust all structural and functional properties of organs. For example, lung concept has structural and functional properties that are not taken into account here.

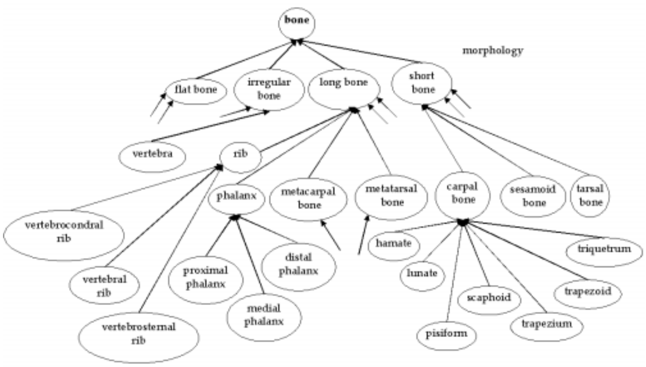

FIG. 3. Taxonomy of the concept bone. Morphology is the criterion used to provide first level classification of bones. Long bones are the bones which have the prevalence of a diameter with respect to the other two. Flat bones have two diameters more relevant than the last one. Short bones have three similar diameters. Irregular bones have no regular shape. This represents a particular view over the properties of bones. For example, rib is classified for its morphology as being a long bone. However, it shares an internal structure similar to flat bones (lack of marrow). This property must be expressed by alternate classification (context).

2.2.2. Is-part-of relation. The part-of relation plays a particular role equally important to that of the ::*is-a-kind-of* relation in the field of semantic networks applied to anatomical knowledge representation. In general, two subtypes of partonomic relations exist that respectively refer to physical and conceptual partition. The first type can be equivalently expressed by the is-consisting-of relation (cell::*is-part-of* tissue, tissue::*is-part-of* organ). The second type is based on the fact that an anatomical entity considered as a whole can be conceptually (arbitrarily) subdivided into two or more parts (organ part::*is-part-of* organ, shaft of femur::*is-part-of* femur, nose::*is-part-of* head, finger::*is-part-of* hand). Equivalently, nose and mouth are parts of the head; larynx, pharynx, and trachea are parts of the neck; bronchus, bronchiole, alveolus, and lung are parts of thorax (topographical view). As an example, Fig. 5 shows the partonomy of the femur.

Moreover, the conceptual subdivision into parts of anatomical structures can be expressed at multiple granularity levels (detail levels). For example, the category long bone can be further divided into (distal head) diaphysis, (proximal head) epiphysis, and shaft. Inheritance will imply that all long bones (like the femur) shall have necessarily these three parts.

Apart from subdivision of anatomical structures into parts, the partonomy relation has been used to express functional membership (functional view). For example, the partonomy of the respiratory system can be accommodated by simply linking the nose, mouth, larynx, trachea, bronchus, bronchiole, alveolus, lung, and pharynx to the involved system.

2.3. *Nonhierarchical Relations*

As shown, the genus and partition properties that concepts exhibit can be represented as hierarchical networks. In contrast, conceptual, physical, spatial, and functional relations that are intrinsically nontransitive have their semantic representation into nonhierarchical tree of concepts. Conceptual relations between categories express an obvious logic arising from the considerations used to build the categories. For example, taking into account that some organs can contain space and recognizing that this feature is a key discriminating property has led to the generation of the category container organ. Necessarily, all container organs will define cavities or, more precisely, organ cavities. This condition can be made formal by explicitly using a conceptual relation such as container organ::*defines* organ cavity where statements organ cavity::*is-a-kind-of* body cavity, body cavity::*is-a-kind-of* body space, and body space::*is-a-kind-of* anatomical space have been established.

This then implies that a consistent definition for organ cavity could be that it is a body space defined by a container

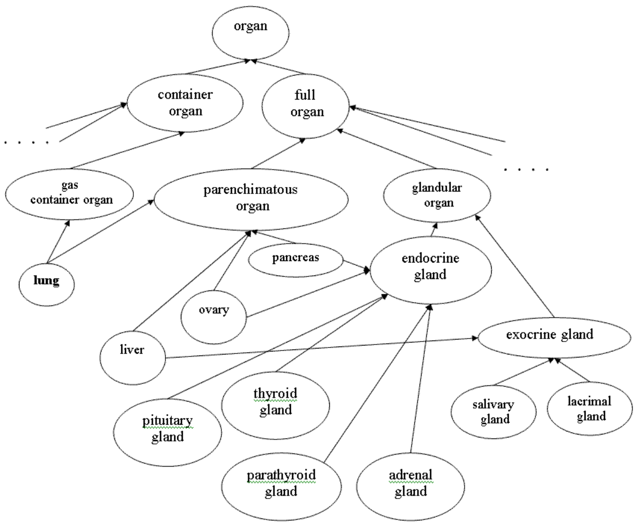

FIG. 4. Polyhierarchies for concept organ. The lung is classified as both a gas container organ and a parenchimatous organ. By applying monotonic inheritance, lung inherits incorrectly characteristics from both container and full organ. For lung concept, inheritance from parenchimatous and full organ is blocked.

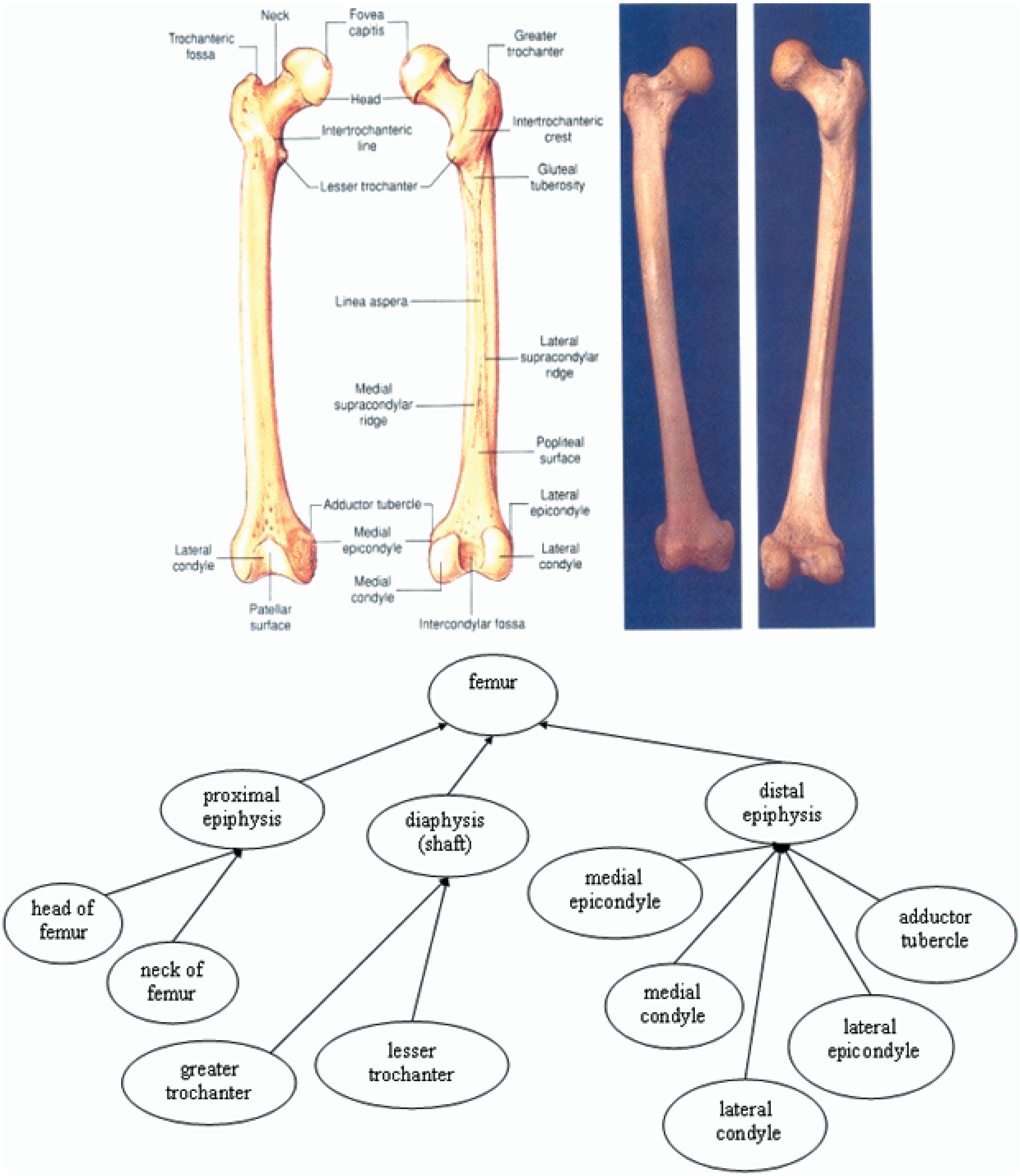

FIG. 5. The partonomy for the concept femur.

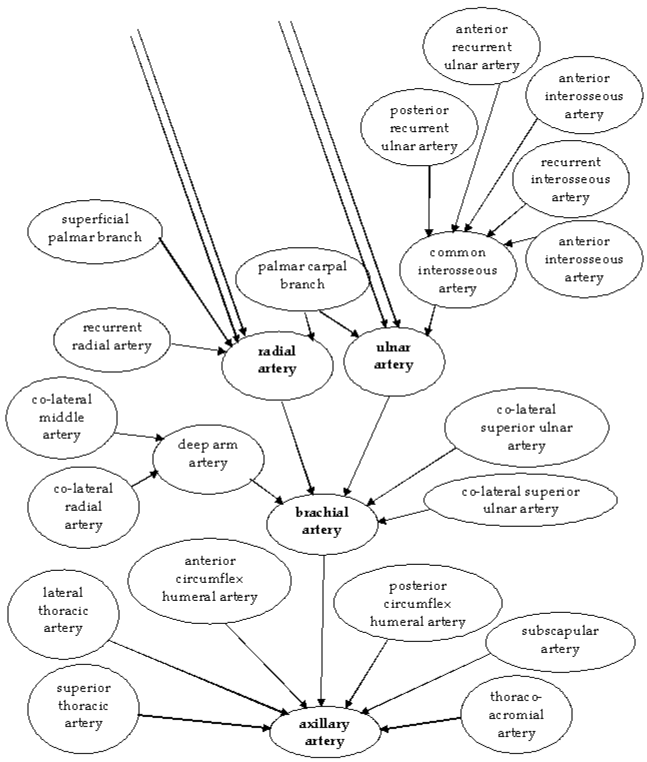

FIG. 6. The *branching-of* relation used to represent the arterial branch of the upper extremity.

organ. As in the case of stomach, it has been classified as container organ and inheritance sanctions that it must be endowed with an organ cavity.

Physical relations are used to represent attributes and common characteristics shared by two anatomical entities. Here the most relevant physical relations we used in the representation are ::*is-connected-to*, ::*is-contained-in*, ::*branching-of*, ::*consisting-of*, ::*interconnecting*, and ::*is-tributary-of*. With the ::*consisting-of* relation (structurally made of, in whole or in part, some physical units, material or matter), the physical properties of anatomical categories and individual structures can be modeled as:

bone::*consisting-of* bone tissue
muscle::*consisting-of* muscle tissue

For the periosteum, endosteum, lamellar bone tissue, and spongiosa, kind-of bone tissues, the inheritance property guarantees that long bone and the other subcategories will be constituted by these kinds of bone tissues. Equivalently, because the heart can be classified as a muscle besides as container organ, it can be assumed that it is constituted by muscle tissue.

The relation of branching between two anatomical concepts can be represented for example as:

artery::*branching-of* artery
femoral artery::*branching-of* external iliac artery
pharyngeal nerve::*branching-of* vagus nerve

The relation *is-tributary-of* can be used to describe a venous tree. For example, the following relations hold:

vein::*is-tributary-of* vein
splenic vein::*is-tributary-of* portal vein

Figure 6 shows a ::*branching-of*-based network for the axillary artery, which is the main arterial branch of the upper extremity. Note that ::*branching-of* relation is not transitive; therefore, the tree in Fig. 6 is not hierarchical. Moreover, that schema does not account for spatiality that must be explicitly expressed by spatial relations, all subsumed by ::*spatially-related-to* relation. These relations refer to properties of adjacency, relative position, path, location, etc. (::*being-adjacent-to, ::connecting, ::entering, ::passing-through, ::being-behind, ::traversing, ::surrounding, ::being-disconnected-from, ::being-externally-connected-to, ::being-partially-overlapped-by*).

In the case of muscles, both origin and insertion into the bones are relevant spatial feature expressed by two relations: muscle::*having-origin-in* bone part and muscle::*having-insertion-in* bone part. For example, psoas major has origin in the transverse processes of the lumbar vertebrae and body of the 12th thoracic vertebra and insertion in the middle surface of the lesser trochanter of the femur. Moreover, both insertion and origin can be further specified. Therefore, the statement "the biceps brachii through its long head has origin at the supraglenoid tubercle of the scapula and through its short head at the coracoid process of the scapula. It has insertion in the tuberosity of the radius" can be represented as muscle part::*having-origin-in* bone part.

Although the origin and the insertion of the muscle determine the movement of the underlying bones, spatial properties do not rely on either function or action. From the parent relation ::*functionally-related-to*, the following set of relations is relevant for anatomical knowledge representation ::*interacting-with, ::innervating, ::supplying-blood-to, ::secreting, ::acting-on* (generic), ::*acting-as-flexor-of, ::acting-as-lateral-rotator-of, ::causing-contraction*, etc.

Figure 7 shows a representation of the muscle of the shoulder and arm where we model partonomic relations along with specific actions and innervations.

A specific functionality of artery consists of supplying

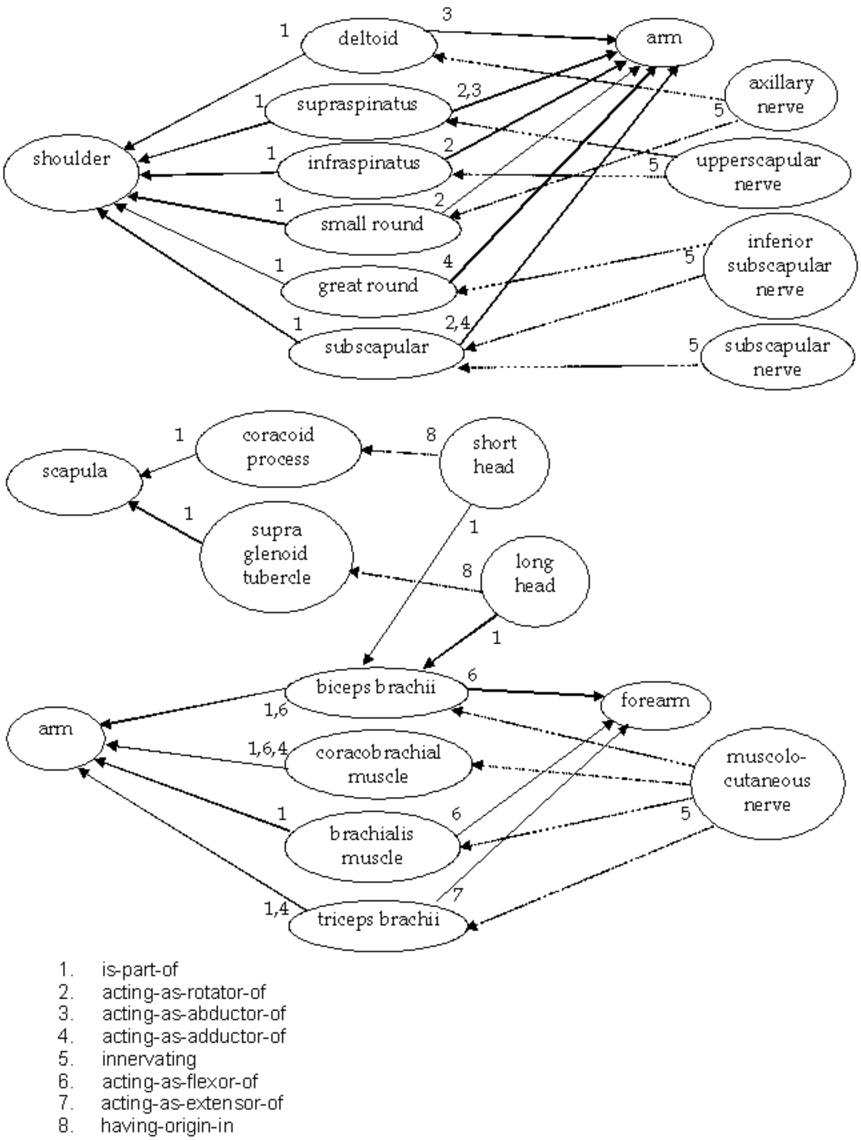

FIG. 7. The muscles of the shoulder and arm: Modeling partonomy, action, and innervations.

blood to tissues of the anatomical structures, whereas nerves provide functional connection between the central nervous system and the body parts. Therefore,

artery::*supplying-blood-to* organ,
renal artery::*supplying-blood-to* kidney,
5th lumbar nerve::*innervating* inferior gemellus (muscle)

are valid relations.

An action is a specific function that an anatomical entity performs on one or more entities such as rising, pulling,

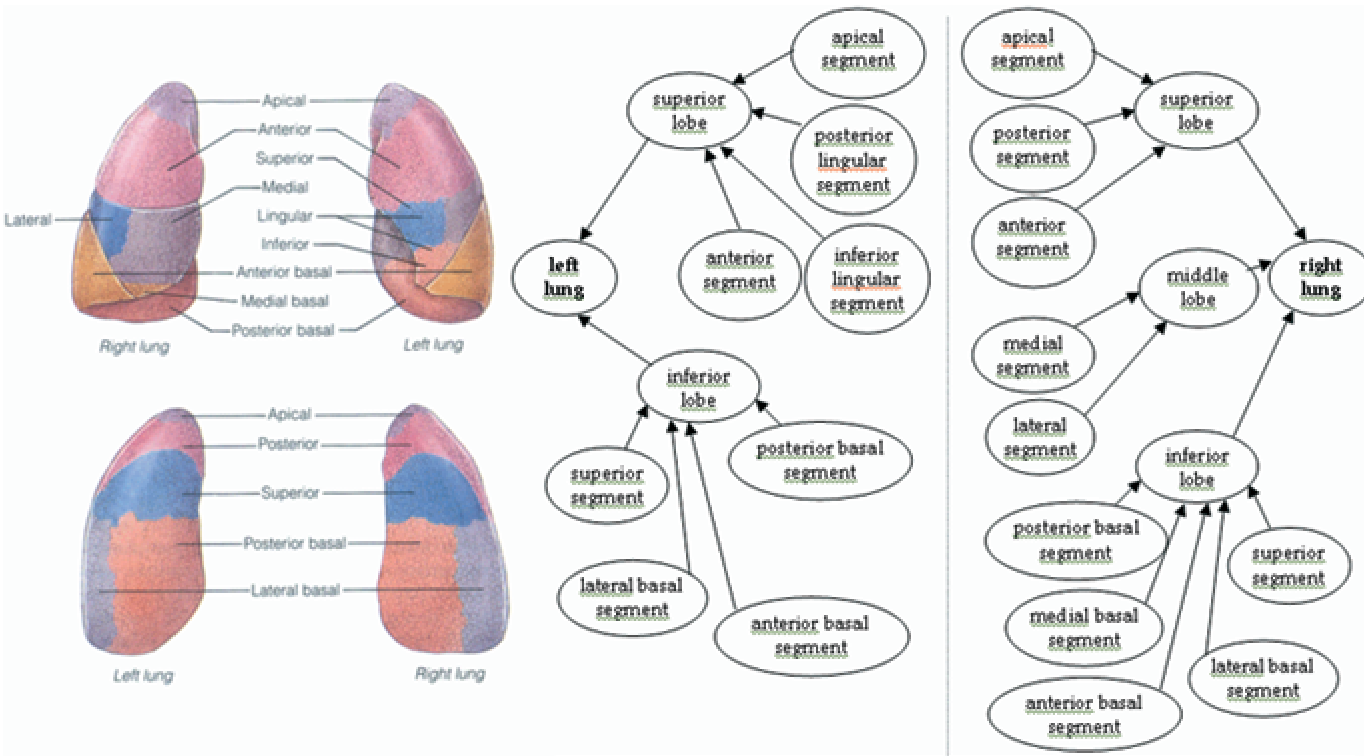

FIG. 8. The partonomy of the lungs. The asymmetry between left and right is evident.

flexing, extending, rotating (medially or laterally), abducting, and adducting. For the muscles, these actions identify their main functionality as in the following example:

muscle::*acting-on* body part
psoas major::*acting-as-flexor-of* thigh,
obturator internus::*acting-as-lateral-rotator-of* thigh.

As an example, four main relations can express muscle actions on the foot as reported in Table 1. Particular attention is to be paid to relationships that involve more than two anatomical objects. For example, representing the case that ventricle and atrium are connected to each other through a valve involves a relation among tree entities that a binary relation cannot accomplish. To include this case we established a new nonhierarchical relationship named association, through which *n* independent concepts are related to one another.

2.4. Knowledge Base Development

Categories (semantic types) define "general" anatomy, whereas the specific anatomical structures (individuals) define "special" anatomy. The relationship between category and individual (physical objects) can be better understood by the following example. Given the category bone, which identifies an element of general anatomy, the category long bone is a specialization of such a category, yet still belonging to general anatomy. In contrast, femur, which is an individual of the category long bone, is an element of special anatomy. Provided that certain nonhierachical and hierarchical relations between categories (general anatomy) and other categories (general anatomy) and between categories (general anatomy) and individuals (special anatomy) are established, the transitive property of hierarchical relations guarantees that the parent attributes are inherited by the offspring. The following example clarifies this point:

::*is-a-kind-of*: long bone (general)::*is-a-kind-of* bone (general)

::*consisting-of*: bone (general)::*is-constituted-by* spongiosa (general)

::*is-part-of*: epiphysis (general)::*is-part-of* long bone (general)

::*is-a-kind-of*: femur (special)::*is-a-kind-of* long bone (general)

TABLE 1
Summary of Muscle Action on the Foot

::acting-as-plantar-flexor-of	*::acting-as-dorsal-flexor-of*	*::acting-as-abductor-with-inversion-of*	*::acting-as-abductor-with-inversion-of*
Tibialis anterior	Peroneus longus	Tibialis anterior	Peroneus tertius
Extensor hallucis longus	Peroneus brevis	Tibialis posterior	Peroneus longus
Extensor digitorum lungus	Gastrocnemius	Extensor hallucis longus	Peroneus brevis
Peroneus tertius	Soleus	Flexor hallucis longus	Extensor digitorum longus
	Plantaris	Flexor digitorum longus	
	Tibialis posterior		
	Flexor hallucis longus		
	Flexor digitorum longus		

::*is-a-kind-of*: femur, left (special)::*is-a-kind-of* femur (special)

::*is-part-of*: femur (special) is-part-of skeleton (special)

From the above, one can automatically infer that the left femur has an epiphysis (it is a long bone) and a spongiosa (as all the bones). If the model is accurately verified, then this approach avoids redundancy caused when relations carrying the same meaning are duplicated across different abstraction levels. This applies also to hierarchical partonomic trees and, for example, if the concept finger is modeled as part of the concept hand and, in turn, hand is part of upper extremity, then there is no need to define finger explicitly as part of upper extremities.

Starting from the above described model and according to point 2 of the first paragraph in the introduction we have built a knowledge base that maps concepts to alphanumeric codes. Then our efforts have been focused on the development of a terminological system to map concepts to language. Toward this aim a terminological system has been constructed according to a compositional strategy. This strategy implies that only *atomic* concepts are explicitly stored in the knowledge base. In particular an *atomic* concept is a concept whose term, used to express it, is constituted by a single item, e.g., bone, organ, muscle, tissue, head. In contrast, nonatomic concepts are expressed by two or more terms. For instance, the concepts full organ, long bone, and striated muscle will be explicitly stored in the knowledge base as an alphanumeric code but the corresponding terms used to express them will be implicitly stored. The term full is an attribute for the concept organ: associating the term full to the term organ we generate the term full organ that is used to reference the concept full organ. In the next paragraphs we will describe how articulated terms like epiphysis of the left femur can be represented and how to connect them to the corresponding concepts.

With this approach, by using atomic concepts and distributed attributes, there is no need to explicitly represent terms that differ in minor ways from each other. If the upper extremity is identified as having left and right laterality, then only arm, forearm, and hand concepts need to be explicitly stored as terms in the knowledge base. By using inference, automated reconstruction can transforms implicit knowledge (the fact that arm, forearm, and hand exist in left and right side) into explicit information. Similarly, the partonomy of the ethmoid bone, which is an unpaired bone in the cranium, constituted by a perpendicular plate and two main lateral masses (left and right) each containing other subcomponents (cribriform plate, lamina orbitalis, middle nasal concha) can be incrementally represented in the knowledge base as reported in Table 2, where explicit and implicit relations have been defined:

The ethmoid bone concept, represented as the composition derived by the anatomical concept bone with ethmoid attribute, is explicitly associated in the partonomy to the composite concept lateral mass having left and right side attributes. Then cribriform plate is explicitly linked to lateral mass so that left and right lateral masses implicitly inherit the part cribriform plate. Equivalently, an explicit link between lamina orbitalis and lateral mass makes possible an implicit link to left and right lateral mass.

In a compositional framework nonhierarchical relations can also benefit from inheritance. Taking into account a functional relation like *acting_as_flexor_of*, the psoas major muscle is a part of the thigh that is a paired body region so that "psoas major::*acting_as_flexor_of* thigh" statement is implicitly duplicated for both right and left psoas major.

Despite its appealing power as a means for knowledge

TABLE 2
Explicit and Implicit Relations Used to Define the Partonomy of the Ethmoid Bone

Concept name	Concept type	Relation name	Relation type	Concept name	Concept type
Ethmoid bone	Composite	::*is-a-kind-of*	Implicit	Bone	Atomic
Lateral mass	Composite	::*is-a-kind-of*	Implicit	Mass	Atomic
Lateral mass	Composite	::*is-part-of*	Explicit	Ethmoid bone	Composite
Left lateral mass	Composite	::*is-part-of*	Implicit	Ethmoid bone	Composite
Right lateral mass	Composite	::*is-part-of*	Implicit	Ethmoid bone	Composite
Cribriform plate	Composite	::*is-part-of*	Explicit	Lateral mass	Composite
Lamina orbitalis	Composite	::*is-part-of*	Explicit	Lateral mass	Composite
Lamina orbitalis	Composite	::*is-part-of*	Implicit	Left lateral mass	Composite

extrapolation, the compositional approach has a sensible limitation: it cannot eliminate completely heuristic knowledge about the domain. In order for composite concepts to be consistently reconstructed, the knowledge base must be provided with specific control operators to prevent the generation of unreal concepts. To make this point clear, let us consider the lungs, the partonomic subdivision of which into lobes and bronchopulmonary segments is depicted in Fig. 8.

It is known that the right lung is morphologically slightly different from the left lung because of the different number of lobes and segments. In fact, the left lung lacks a central subdivision (middle lobe) that is present in the right lung. The compositional representation of this information content takes into account the concepts lung, lobe, and segment, whereas the laterality attribute (left, right, anterior, superior, medial, etc.) is compositionally represented, as will be clarified below. However, a nonsupervised reconstruction is not able to account for heuristic facts such as lung asymmetry and would generate erroneous concepts. By assuming that the concept lung exists in left and right instances and is divided into three different pulmonary lobes (superior, middle, and inferior), the reconstruction engine would assign all three lobes to both right and left lungs erroneously. This implies the need for reconstruction to be driven by a heuristic rule breaking the link between left lung and middle lobe.

As a consequence, inheritance in a knowledge base developed through a compositional approach has two separate features: one related to semantic level as described in Section 2.2.1 (no matter how concepts are stored) and the other related to how concepts have been represented through terms. An enumerative strategy avoids this last problem by explicitly storing composite concepts through corresponding defining terms: the concept head of the left femur will have an entry in a concept structure through a concept code and an entry in the term structure containing the string “head of the left femur.” In contrast, in our compositional strategy only the code of that concept has been explicitly stored in a data structure that we have named *Composition*, whereas the corresponding term has been compositionally represented. A devoted structure in the knowledge base has been designed to implicitly define composite concepts.

Let us clarify this process by an example. Head and neck are terms that are separately coded with respect to concepts that can be indicated through these terms. For example, concept “head” as a body part will have a concept code in a data structure named *BasicAnatomicalConcept* that will be linked to other concept codes to define its semantics. In addition, it will be associated to a term code with the string “head.” In contrast, the concept “head of the proximal epiphysis of the left femur,” being a composite concept, will have a code in the *Composition* data structure. Similarly, epiphysis is an anatomical concept and part of femur and proximal epiphysis is one of its composite concepts. Table 3 shows how the concept “head of the proximal epiphysis of left femur” has been represented in our compositional framework.

The top of Table 3 shows the composition of the concept as a join between a root concept and an attribute. An attribute can be either a feature or concept For example, the term that refers to the concept left femur has been built by composing the term femur and the laterality feature left. At the bottom of Table 3, we report the related explicit and implicit semantics. In particular, this development strategy implies that the concept “head of proximal epiphysis of left femur” has been explicitly stored as an alphanumeric code, whereas the corresponding term has not. The following steps are required to iteratively reconstruct that term:

(1) Given the concept code, the procedure recovers the kernel concept code (head of proximal epiphysis of femur), that is in turn a composite concept (first row);

(2) The concept code of head of proximal epiphysis of

TABLE 3
Terminological Composition of Concept "Head of Proximal Epiphysis of Left Femur"

Terminology					
Concept code	Concept type	Kernel concept code	Kernel concept type	Attribute code	Attribute type
Head of proximal epiphysis of left femur	Composite	Head of proximal epiphysis of femur	Composite	Left	Feature
Head of proximal epiphysis of femur	Composite	Head of proximal epiphysis	Composite	Femur	Concept
Head of proximal epiphysis	Composite	Head of epiphysis	Composite	Proximal	Feature
Head of epiphysis	Composite	Head	Atomic	Epiphysis	Concept
Left femur	Composite	Femur	Atomic	Left	Feature
proximal epiphysis	Composite	Epiphysis	Atomic	Proximal	Feature
Epiphysis of femur	Composite	Epiphysis	Atomic	*Femur*	Concept
Proximal epiphysis of femur	Composite	Proximal epiphysis	Composite	*Femur*	Concept

Semantics

Explicit
Epiphysis of femur is part of femur
Head of epiphysis of femur is part of epiphysis of femur
Implicit
Left femur is-a-kind-of femur
proximal epiphysis of femur is part of femur
proximal epiphysis of left femur is part of left femur
Head of proximal epiphysis of femur is part of proximal epiphysis of femur
Head of proximal epiphysis of left femur is part of proximal epiphysis of left femur
Head of proximal epiphysis of femur is part of femur
(Intermediate level in the hierarchy has been hidden
Head of proximal epiphysis of left femur is part of left femur
(Intermediate level in the hierarchy has been hidden

Note. Concept composition implies implicitly genus relation between root composite concept and concept (left femur is-a-kind-of femur). In the knowledge base implementation we have constrained this property by a devoted flag (see Table 4) in *Composition* data structure.

femur is retrieved (second row) as composition of the above concept with femur attribute;

(3) The root concept head of proximal epiphysis is retrieved (third row) as the composition of the above concept with proximal attribute;

(4) Equivalently, the root concept head of epiphysis is retrieved (fourth row);

(5) The anatomical concept head is retrieved (fifth row);

(6) The reconstruction engine generates the needed term head of proximal epiphysis of left femur.

The relevant semantics are reconstructed by using inheritance from the implicit representation as depicted at the bottom of Table 3. In particular, the semantics can be reconstructed according to the following steps:

(a) Given that head of epiphysis of femur::*is-part-of* epiphysis of femur (see bottom of Table 3) and proximal epiphysis is compositionally a kind of epiphysis (see top of Table 3), the relation head of proximal epiphysis of femur::*is-part-of* proximal epiphysis of femur holds;

(b) Given that head of proximal epiphysis of left femur is compositionally a kind of head of proximal epiphysis of femur, the relation head of proximal epiphysis of left femur::*is-part-of* proximal epiphysis of left femur holds.

The knowledge base has been implemented by utilizing a relational model and stored in a database named *AnatomyKnowledge*. The first data structures we developed in *AnatomyKnowledge* were *BasicAnatomicalConcept*, *Definition*, and *Term* table, the latter containing only anatomical terms such as organ, bone, long, and short without any explicit conceptual reference. Moreover, we have classified distinct terms as either preferred or synonym. The *BasicAnatomicalConcept* table (see Table 4 and Fig. 9) has been used to link anatomical concepts to terms and definitions (organ, bone). Nonatomic concepts (composite) like long bone, short bone, metatarsal bone, and metacarpal bone have been implicitly stored as a link through two separate data structures: the table *Attribute* containing a term code corresponding to the attribute value (long, short, . . . , red, green, . . . , lateral,

TABLE 4
The Main Relational Tables of the *AnatomyKnowledge* Database

BasicAnatomicalConcept table

ConceptCode	*TermCode*	*DefinitionCode*
C000000128	T000002342 (bone)	D00ewf43r3
C000008221	T0000r4823 (hand)	D0045187g4

Attribute table

AttributeCode	*Attribute_TypeCode*	*Attribute_ValueCode*
A01a	(dimension)	(long)
A01b	(shape)	(short)
…	…	…
B01a	(laterality)	(left)
B01b	(laterality)	(right)
C01a	(laterality)	(superior)
C01b	(laterality)	(anterior)
C01c	(laterality)	(lateral)
C01a	(laterality)	(superior)
..	..	..

Composition table

Composite ConceptCode	*RootConcept Code*	*AttributeCode*	*Semantic*
..	..	..	..
CC00003212 (long bone)	C000000128 (bone)	A01a	TRUE
CC00002821 (left hand)	C000008221 (hand)	B01a	TRUE
		..	..

Network table

NetworkID	*ConceptCode*	*Parent ConceptCode*	*RelationCode*
N00423e001	C000000128 (bone)	C000000006 (organ)	r01a (is-a-kind-of)
N004f34002	C000002314 (artery)	C000000349 (vessel)	r01a (is-a-kind-of)
N000w21367	CC000s0453 (systemic artery)	C000002314 (artery)	r01a (is-a-kind-of)
…			
N3w3128897	C000001231 femur	CC00003212 (long bone)	r01a (is-a-kind-of)
…	…	…	…

medial, . . .) and a specification term code corresponding to attribute types such as size, dimension, color, shape, density, and the table *Composition* storing composite concept codes. Semantic relations among concepts have been distributed across two different data structures. In the *Composition* table the field *Semantics* (Boolean value) has been constructed to specify that the composite concept assumes semantic distinction from the parent concept (see Table 4). In Table 4 *Network* two generic concepts are explicitly linked by a relation with an associated context.

Figure 9 shows the relational schema that we designed to implement the knowledge base. In particular, note the data structure we named *CrossMapping*. In this table we matched the codes of the concepts used with the corresponding UMLS concept codes. This allows the reconstruction stage to automatically obtain from the UMLS server (http://umlsks.nlm.nih.gov/) the concept information taken from major digital biomedical terminological sources.

2.6. *Information Reconstruction and Inheritance*

To retrieve understandable information from the knowledge base the following main issues have been taken into account: (a) Most terminological knowledge has been implicitly stored (compositional approach), as a result of which a reconstruction procedure is required; (b) semantics of concepts can be retrieved by applying inheritance, but must be carefully verified through a supervision based on stored heuristic knowledge; (c) while being easily manageable, the relational data model used to implement the knowledge base and SQL-based database query engine do not provide either adequate data structures or methods to efficiently retrieve stored information; (d) information must be presented to the user in a suitable way.

In order to address these problems and fit the requirements, we conceived and developed our system in the following way:

The application has been subdivided into two parts: a server side and a client side.

The server side performs the following actions:

Receives user requests from client interface
Builds queries
Runs queries on the databases
Reconstructs information
Sends result to client interface (data and images)

The client side has the following functionalities:

Data display
Managing user requests
The server side collects data from two database systems
AnatomicalKnowledge relational database
ImageMap database

The *ImageMap* database, accounting for coded image data, has been used to verify the usability of the developed knowledge base.

The server-side software framework has been developed according to the object oriented paradigm and implemented in the Java language, using a set of classes as *TermClass*, *ConceptClass*, *DescriptionClass*, *RelationClass*, *ConceptTreeClass*, *ConceptTreeArrayClass*, each characterized by a set of operators (see Appendix B) that generates a safe SQL statement, issues a corresponding query to the *AnatomyKnowledge* database, and formats the results for the defined data structures. For example, *GetAscendingTree*, a method of *ConceptTreeArrayClass* class, acts as a reconstruction operator by receiving as input a *TermClass* object corresponding to an anatomical concept, a *RelationClass* object specifying the involved relation, and if necessary a user-needed depth level, and giving back as output an array of the parent concept trees. Similarly, *GetDescendingTree* gives back as output an offspring concept tree. By opportunely composing the operators, the reconstruction engine is able to explicitly obtain a wide range of anatomical information in the form of constrained views. As mentioned above, the key point is that the reconstruction engine, automatically applying inheritance (inference), cannot ignore the possibility of generating unreal concepts. Let us clarify this issue by describing the reconstruction procedure for the partonomy of the lung and left lung concepts based on the *GetDescendingTree* operator. In generating the lung partonomic tree, the reconstruction engine automatically recovers the abstraction level of the corresponding concept and the corresponding detail level. In this case the concept lung has been explicitly linked (*Network* table) to the lobe concept and lobe to segment, and the reconstruction engine directly builds the lung-lobe-segment tree. The concept left lung, while being a specialization of the concept lung, has been represented at a refined detail level, although not explicitly stored in the *Network* table. Initially the reconstruction engine looks for the first parent (compositionally represented) the partonomy of which has been explicitly expressed (lung, in this case). Then the system takes into account the concept lobe as the first level part of the concept lung, attempting to associate each specific lobe (superior, inferior, and middle laterality from the *Composition* table) to left lung, applying inheritance. The operator *VerifyFacts* (see the Appendix), which implements a consistency rule based on heuristic knowledge between the two matched concepts, excludes the generation of possible unreal concepts. Heuristic knowledge has been stored in a data structure named *Facts* in which each entry indicates an explicit block of inheritance between two concepts (see Table 5). In particular, by finding an entry in the *Facts* table that links left lung to middle lobe, inference avoids the generation of the concept middle lobe of the left lung. By moving to segments (apical, lingular, basal, posterior, anterior, medial, etc.) the inference engine repeats the control, excluding, for example, medial segment from superior lobe (see the Appendix for the pseudo code).

TABLE 5
The *Facts* Table Used to Store Heuristic Knowledge

ConceptCodeA	*ConceptCodeB*	*RelationCode*
…	…	…
CC00rs000 (lung)	CC0000023 (full organ)	(is-kind-of)
…	…	…
CC00rs001 (left lung)	CCfe342112 (medial segment)	(is-part-of)
CC00rs001 (left lung)	CC00asb22 (middle lobe)	(is-part-of)
…	…	…

Note. For the entries the inheritance is to be blocked.

Each time a property must be subsumed from one concept to another, the *VerifyFacts* operator verifies from the *Facts* table whether a possible block exists. This corresponds to assuming that inference is automatically enabled apart from explicit coded blocking.

Based on this information reconstruction procedure, the developed software system provides the ability to issue the following requests:

- All anatomical concepts that contain or exactly match a user-defined key word;
- All the synonyms given a preferred term;
- Progressive detailed representations for a concept;
- All anatomical concepts that belong to a specified body region and/or to a specified body system;
- The hierarchical tree (taxonomic or partonomic—ascending and/or descending) for a specified concept in a specified context;
- The network of characteristics related to a concept (non-hierarchical relationships);
- Constrained views: all anatomical objects involved in a certain function, located in a specific body region, or satisfying other particular conditions.

3. USE OF THE DEVELOPED KNOWLEDGE BASE

The utility of the developed anatomical knowledge base has been demonstrated by assembling a prototypical client side interface. In addition to knowledge-base query, the client side has been conceived to provide symbolic access to spatial information, constituted by enhanced features like segmented pixel regions or contours, of a reduced set of the images of the Visible Human Dataset (VHD) [16]. The appealing idea here is that of enabling the user to interactively construct anatomical statements to recover both symbolic and spatial content. The latter consists of partially segmented images in the abdomen from the VHD contained in a test set provided by Gold Standard Multimedia. The image segmentation has been refined to cope with the detail level expressed in the *AnatomicalKnowledge* database. In particular, each labeled pixel has been assigned to the corresponding anatomical entity (concept code) with the highest level of granularity. For example, pixels of the region imaging the left lung will never be assigned to the left lung concept—that would cause information loss—but rather will be assigned to a composite concept like anterior basal segment of the inferior lobe of the left lung. Therefore, at run time when such pixels will be picked, the corresponding concept and its properties can be reconstructed with any (even user defined) granularity level.

Spatial information has been structured in such a way as to take into account: (a) pixel membership, (b) contours of anatomical structures, and (c) relations between labeled pixels and *AnatomicalKnowledge* concept codes. With respect to the first two points, two types of files have been generated for each image: a mask file in which each value, corresponding to an image pixel, has been assigned by a label to an anatomical structure, and a contour file, containing pixel coordinate pair lists each corresponding to a structure contour (see Fig. 10). The relation between labeled pixels and *AnatomicalKnowledge* concept codes has been set up through a database named *ImageMap* in which labels are externally linked to concept codes. The *ImageMap* database also takes into account the fact that an anatomical structure can appear in multiple images as an image region as well as a contour.

The overall software system has been conceived as a two-layer framework: the client side that provides visualization and user-interface facilities and the server side that receives user input via RMI (remote method invocation) technology, processes user requests, queries the databases through the reconstruction engine, and delivers results to the client side. The client-side application, consisting of a set of panels, provides the user with the ability to visually formulate constrained queries to the image database *ImageMap* via different user interaction modes.

Figure 11 shows the operation pipeline for retrieving semantic information about an anatomical structure corresponding to a pixel picked by the user. First, the server-side engine, after receiving the picked pixel coordinates and the index of the currently displayed image, looks in the corresponding mask file for the label of the related anatomical structure. Then it opens the corresponding contour file to find out the contour (pixel coordinate pair list) of the involved anatomical structure and from the *ImageMap* database it obtains the anatomical structure code mapped into the *AnatomyKnowledge* database. Then, it queries the *AnatomyKnowledge* database to retrieve general information about the involved anatomical concept. All these results are then delivered to the client side where the contour is depicted as superimposed on the image and concept information is visualized via the interface.

Some different modes are foreseen for querying the knowledge base. A query by term represents the standard search by key words: the user can recover the list of terms that exactly match, contain, or are synonyms of the required keyword. As shown in Fig. 12, the result consists of a list of terms that can be alternatively selected to recover the basic semantic information about the corresponding concept (definition and taxonomy). Semantic queries provide access to anatomical concepts that fit a semantic constraint defined by a user-specified relationship. The result consists of tree-like views in which each single item can be picked to show relevant information.

A visual browsing panel provides the user with the ability to retrieve image content corresponding to a selected item in a tree of concepts obtained after a semantic query (see Fig. 13). In this case, the user has selected the item "left

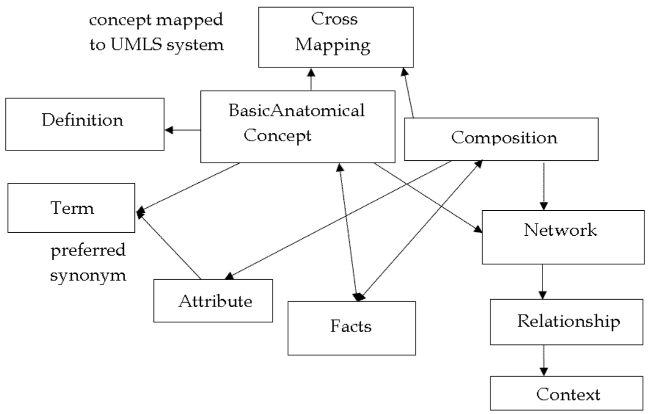

FIG. 9. Relational model of the knowledge base.

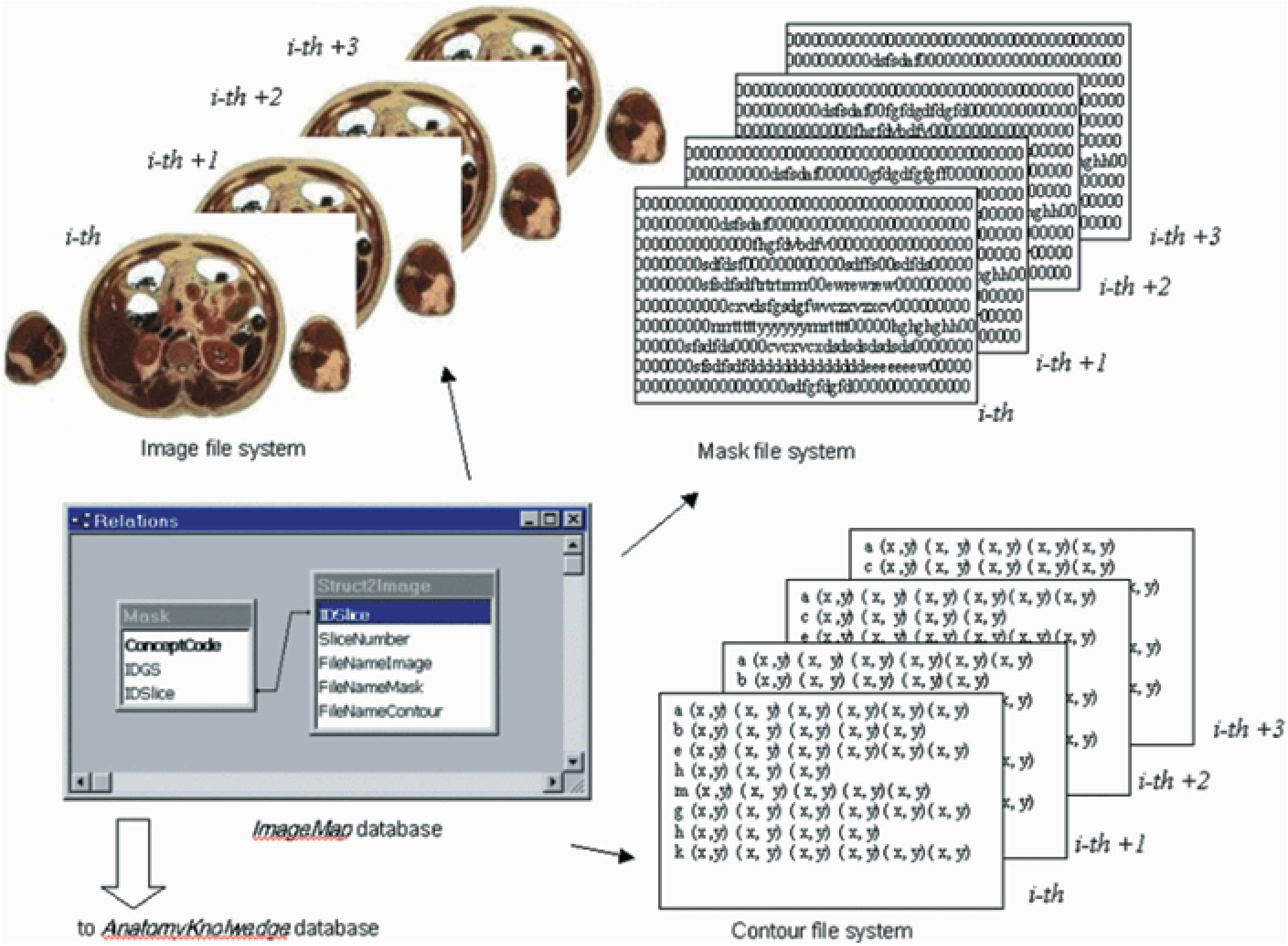

FIG. 10. The relational schema of the *ImageMap* database in connection with the visual data sources.

kidney." This action corresponds to retrieving all of the VHD images that contain that anatomical structure. In this case the result is many-fold: in the left box, a list of the retrieved images is displayed; in the central box, containing a sagittal image of the whole body reconstructed from axial original images, a band (yellow lines) is displayed, superimposed on the image, which indicates the location in the dataset of the retrieved images. The red line accounts for the currently displayed image; in the right box, a representative image of the retrieved image set is visualized at a user-defined resolution (50% in the example of Fig. 13) with a superimposed contour (in green) of the structure for which the user was looking.

Then by right clicking in the proximity of the structure contour, the main semantic information is listed in a floating box. Alternatively, the user can pick (via the left mouse button) any other pixel in the image in the right box to retrieve information about the corresponding anatomical structure.

Figure 14 shows a Constrained query panel in which the user can construct an arbitrary (yet, system-driven) query (in this case "select all muscle in the arm") and visualize the results. Then the user can click with the mouse close to the contour of any structure ("pronator teres, right" in this case) to retrieve features about the corresponding anatomical structure listed in a floating box.

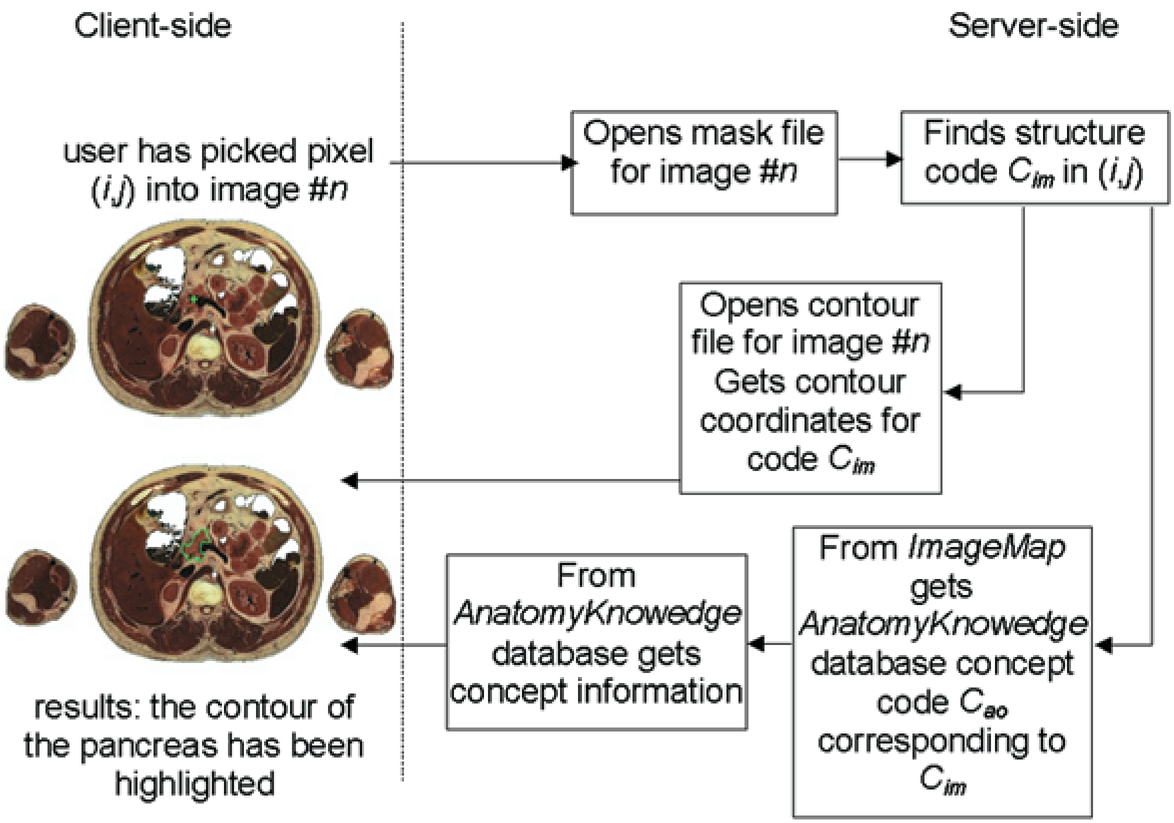

FIG. 11. Operation pipeline to retrieve information after the user has picked a pixel into the image.

4. DISCUSSION AND CONCLUSION

This paper constitutes a first-stage attempt to consistently represent anatomy in a knowledge base. The main novelty of our development consists of the ability to reproduce heterogeneous views over anatomy by modeling genus and partitive properties through hierarchical relations and physical, spatial, and functional attributes through nonhierarchical relations. In particular, the work has focused attention on the problem of modeling anatomical concepts according to a multiple-classification paradigm. Toward this aim, both constrained polyhierarchies and context-dependent strict hierarchies have been used. In addition, the explicit representation of partonomic relations has allowed accommodation of systemic and topographical views in a unique framework. The power of the representation has also been enriched by defining several levels of granularity for concepts in terms of abstraction and detail levels as discussed above. This has allowed accommodation of the user need for manipulating the same concept at different levels of specification. Attention has been paid to the inheritance feature by implementing a reconstruction engine that uses heuristic knowledge to explicitly block wrong inheritance.

We stressed, moreover, the role of the development strategy we adopted (compositional approach) by describing the advantages and weaknesses. The compositional strategy used to organize concepts has resulted in the following advantages:

- Definition of the main semantics directly at the conceptual level of the knowledge base disregarding specific terms;
- Prevention of explicitly representing huge amounts of redundant information in contrast to an enumerative strategy;
- Ability to store both semantic and nonsemantic attributes in the knowledge base.

However, an ad-hoc supervisor to reconstruct composite concepts is required. Some 3000 anatomical concepts have been inserted in the *AnatomicalKnowledge* database along with 1500 relations.

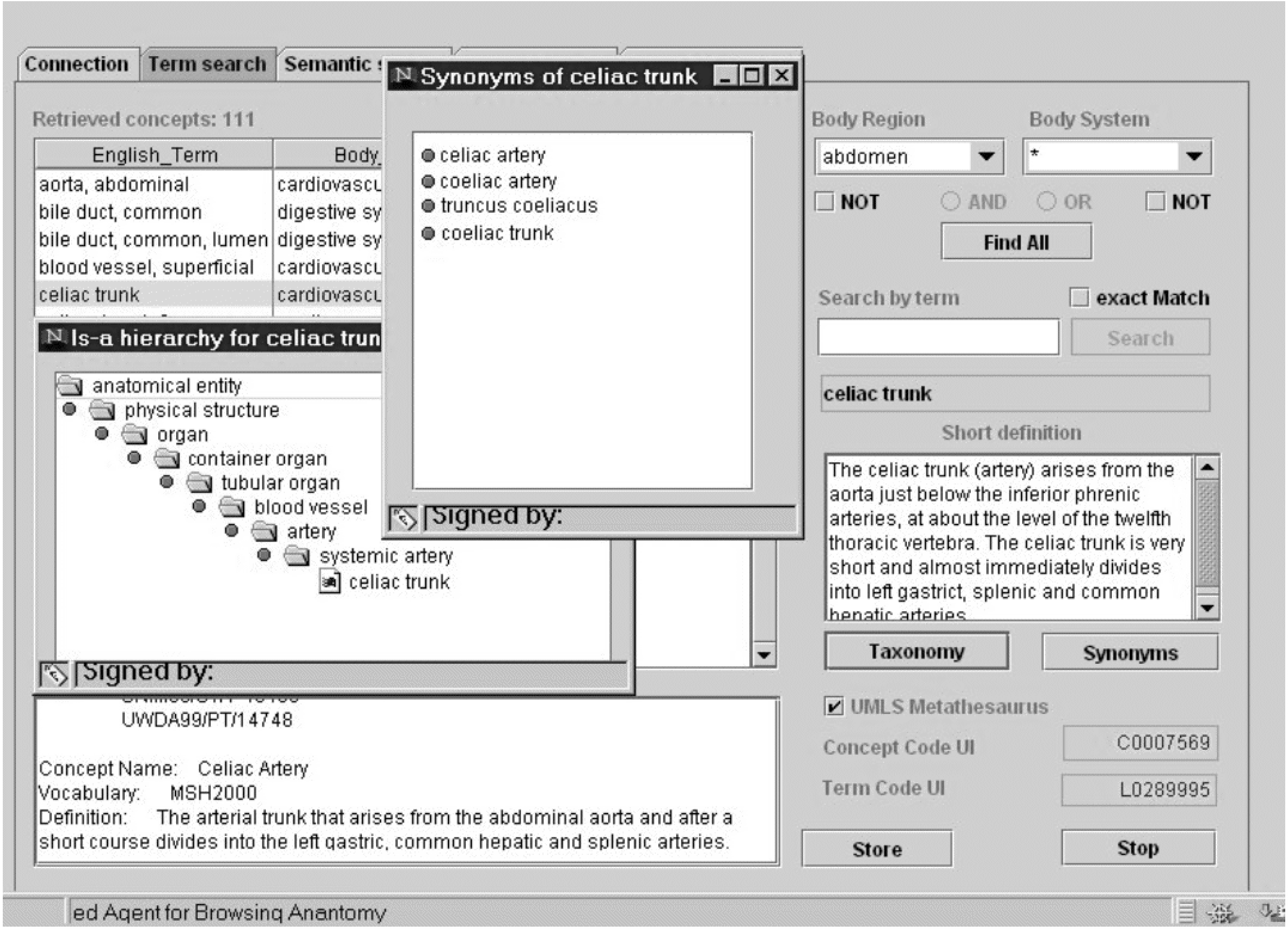

FIG. 12. The "term search" panel. The user has been looking for anatomical concepts into the abdomen. The results was a term list (upper left box). Then he/she selected an item (celiac trunk) of the list and the definition appeared in the right centered box. Then synonyms and taxonomy have been retrieved (floating boxes). Moreover, in the lower left box additional concept information has been listed which has been automatically retrieved from UMLS server.

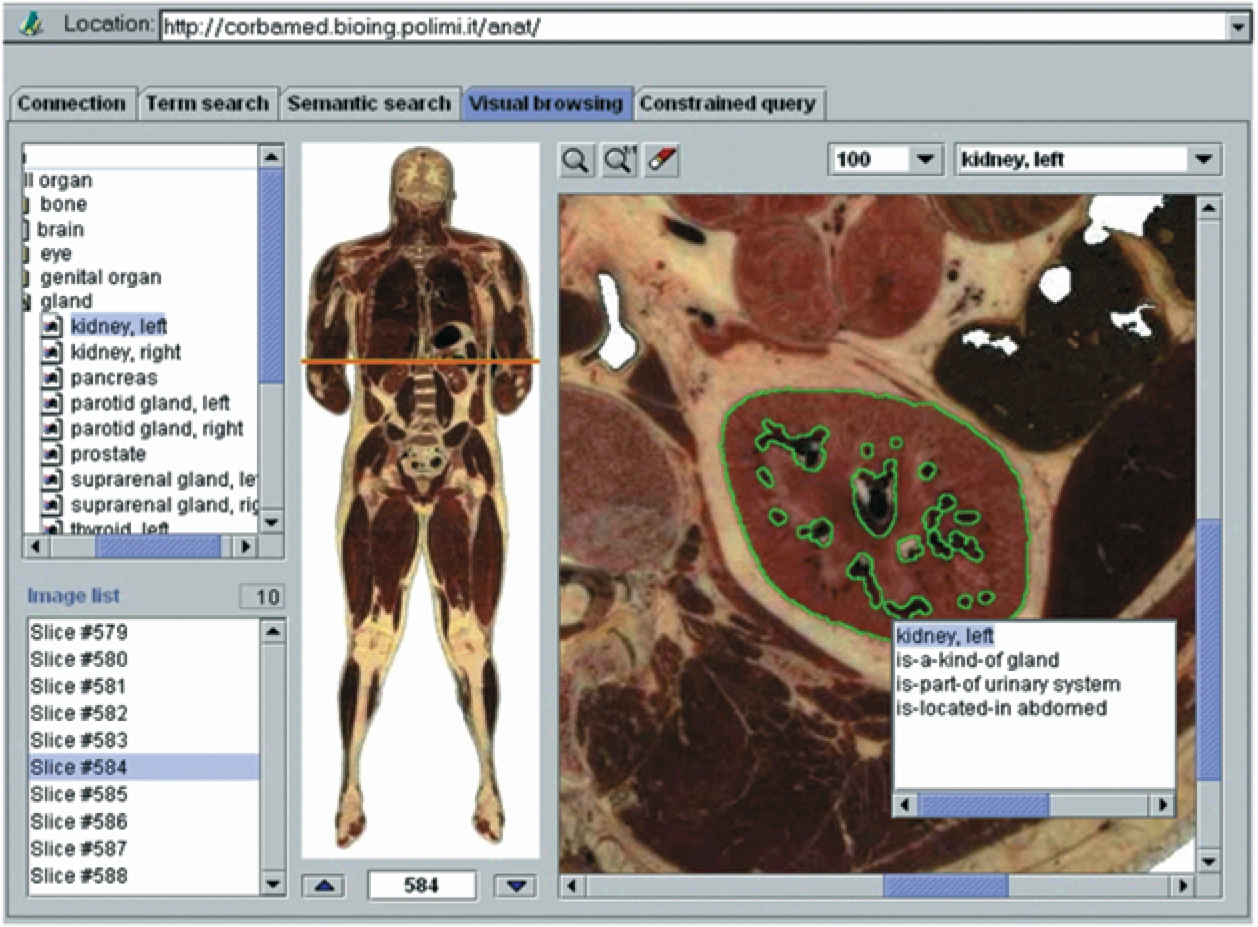

FIG. 13. The "visual browsing" panel. After a semantic hierarchical query performed into "semantic search" panel, the user can select an item (left kidney) to look for the VHD images where the corresponding structure is imaged. Then the user has picked (right mouse button) nearby the contour to show main semantic information (floating box).

With regard to anatomy, we oriented our modeling toward beginners, so we focused our conceptualization on some organs, muscles, bones, nerves, and vessels. However, we acknowledge that other organs and particular anatomical entities cannot be so easily and consistently represented.

The initial tests performed have demonstrated the reliability and consistency of the developed knowledge base, although the concept category should be systematically validated with the aid of anatomists for the knowledge base to constitute the kernel of a representation for clinical concepts [20]. Articulated concepts such as "fracture of the left femur" can be easily composed by linking the composite anatomical concept left femur with the anatomical clinic concept of fracture in the representation.

Similar to other recent efforts [21, 22] aimed at integrating spatial and symbolic information, we developed a prototypical user interface to the knowledge base which allows one to semantically navigate an anatomical image set (2D) derived from the Visible Human Dataset organized into a framework composed by a database (*ImageMap*) and two file systems. Although this software system is in the first stage of development and no systematic evaluation has been carried out to date, it has been used to demonstrate the reliability and consistency of the developed knowledge base. Moreover, it has the great advantage of providing the ability to access (constrained views of) image content through arbitrary user-generated statements (e.g., retrieve all muscles in the left arm) for which typical results are shown by Figs. 13 and 14. As a limitation, the visual user interface lacks three-dimensional visualization. We recognize this as a funda-

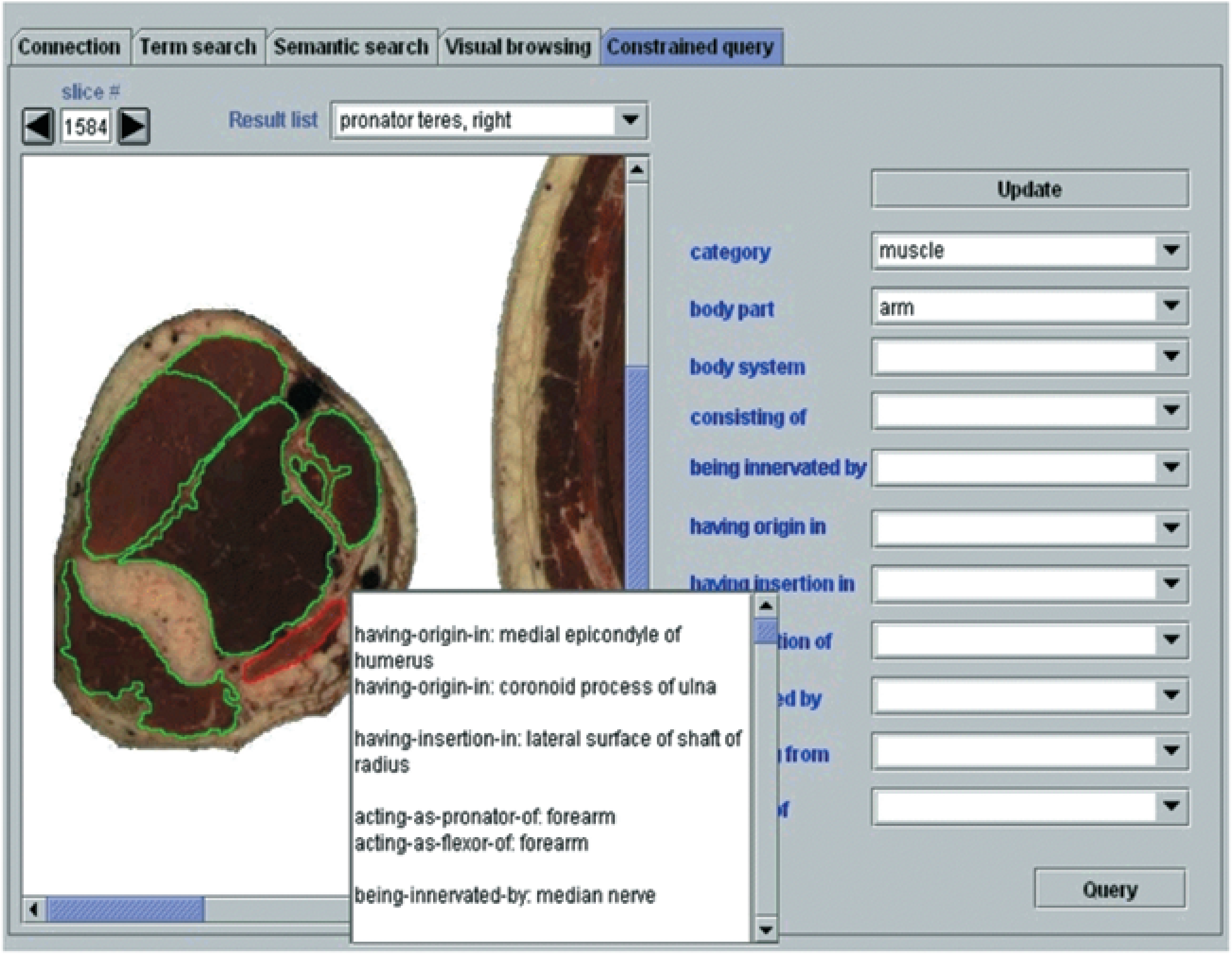

FIG. 14. The "constrained query" panel. The user can build a semantic query by incrementally selecting constraints (query by example) in the right panel side. In this case, the constrained query is: "select all muscle in the arm." Then the user has picked with the mouse close to the contour of structure "pronator teres, right." In this case, a query is submitted to retrieve semantic characteristics of the involved structure (floating box).

mental requirement for such types of applications. Therefore, future developments of the proposed work will include improvement of the anatomical concept organization and introduction of 3D spatial information of the indexed anatomical structures.

In conclusion, we believe that the proposed work represents a suitable contribution in the field of medical informatics in terms of computer-based knowledge representation and semantic access to images.

APPENDIX

(A)

Here we report the definitions of the two major relations (obtained from UMLS server: http://umlsks3.nlm.nih.gov)

::*is-a-kind-of*: This is the basic hierarchical link in the semantic network. If a concept “is-a-kind-of” another concept, then the first concept is more specific in meaning than the second concept.

::*is-part-of*: Composes, with one or more other physical units, some larger whole. This includes component of, division of, portion of, fragment of, section of, and layer of.

The definition of the other relations does not differ from standard definition.

(B)

The following pseudo code shows the main classes and operators devoted to information reconstruction:

```
ConceptClass::GetParents(RelationClass relationObj)
{
        string sqlCode = SELECT ParentConceptCode FROM Network
                                WHERE ConceptCode = this.code AND
                                      RelationCode = relationObj.code
                         UNION SELECT ConceptCode FROM Composition
                                WHERE CompositeConceptCode = this.code AND
                                      Semantics = TRUE
        result = RunQuery(sqlCode)
}
```

Given a relation object it builds the SQL query to recover the parents of the *ConceptClass* and issues it.

```
ConceptClass ConceptClass::GetNextParent( )
{
        if result != NULL
                  ConceptClass parentConcept;
                  parentConcept.code = result.GetNextValue( );
                  parentConcept.GetInfo( );
                  return parentConcept
        else
                  return NULL
}
```

After issuing the query *GetParents* it retrieves the parent object.

```
ConceptClass::GetConceptfromTerm(TermClass termObj)
{
        for i = 1 : termObj.nltems
                 if ( IsConcept(termObj.ParseTerm(i), this.code ))
                        GetInfo( );
                        return 0
        return 1
}
```

Given a TermClass object find, if existing, the corresponding *ConceptClass* object.

```
ConceptTreeClass::GetAncestors(ConceptClass conceptObj, RelationClass
relationObj, integer currLevel, integer finalLevel)
{
     ConceptClass conceptObjParent;
     conceptObj.GetParents(relationObj)

    while ( (conceptObjFather=conceptObj.GetNextParent (j)) !=NULL )
          AddFather(conceptObj, conceptObjFather)
          if (currLevel <= finalLevel)
                   GetAncestors(conceptObjFather, relationObj, ++ currLevel,
                   finalLevel)
    }
}
```

It is a recursive method of *ConceptTreeClass* that attaches to the tree object the tree built with the parents.

```
ConceptTreeArray::GetAscendingTree(TermClass termObj, RelationClass
relationObj, integer finalLevel)
{
        integerlevel = 0
        ConceptClass conceptObj
        if (conceptObj).GetConceptfromTerm(termObj)!=0
                  END
        for all contexts of relationObj
        {
                   ConceptTreeClass conceptTree
                   ConceptTree.GetAncestors(conceptObj, relationObj, level,
                   finalLevel)
                   AddTree(conceptTree)
        }
        return conceptTreeArray
}
```

It reconstructs all the ancestor concept trees (one for each context of the involve relation) for a given
ConceptClass object derived by a *TermClass* object and a *RelationClass* object.

```
ConceptClass::GetSons(RelationClass relationObj)
{
        string sqlCode;
        if (relationObj.name==is-a-kind-of)
                   sqlCode = SELECT ConceptCode FROM Network
                                   WHERE ParentConceptCode = this.code AND
                                         RelationCode = relationObj.code
                             UNION SELECT CompositeConceptCode FROM Composition
                                   WHERE ConceptCode = this.code AND Semantics = TRUE
        //Run query
               this.result = RunQuery(sqlCode)

        else if (relationObj.name==is-part-of)
                    if (this.isAtomic==TRUE)
```

```
                sqlCode = SELECT ConceptCode FROM Network
                WHERE ParentConceptCode = this.code AND
                     RelationCode = relationObj.code
                //Run query
                        this.result = RunQuery(sqlCode)
                else
                        ConceptClass atomicConceptObj
                        atomicConceptObj = this.GetAtomicConcept( )
                sqlCode = SELECT ConceptCode FROM Network
                           WHERE ParentConceptCode = atomicConceptObj.code AND
                                RelationCode = relationObj.code
                //Run query
                        this.result = RunQuery(sqlCode)
}
According to the kind of involved relation and the typology of the concept
(anatomical/nonanatomical) it builds the corresponding query to retrieves
the children and issues it
ConceptClass ConceptClass::GetNextSon( )
{
      if result != NULL
               ConceptClass sonConcept;
               sonConcept.code = result.GetNextValue( );
               sonConcept.GetInfo( );
               return sonConcept
      else
               return NULL
}
After issuing the query GetSons it retrieves the son object.
ConceptTreeClass::GetOffspring(ConceptClass conceptObj RelationClass
relationObj, integer currLevel, integer finalLevel)
)
{
      ConceptClass conceptObjSon

      conceptObj.GetSons(relationObj)
      while ((conceptObjSon=conceptObj.GetNextSon( ))!=NULL)
      {
            if ((conceptObj.isAtomic( )==FALSE) AND (relationObj.name!=
            is-a-kind-of))
            {
            conceptObjSon.GetSons(RelationClass(is-a-kind-of, code))
            while ((conceptObjSon = conceptObj.GetNextSon( ))! = NULL)
      {
            if (VerifyFacts(conceptObj, conceptObjSon) == FALSE)
      {
            AddSon(conceptObj, conceptObjSon)
            if (currLevel <= finalLevel)
            GetOffspring(conceptObjSon, relationObj, ++currLevel, finalLevel)
      }
```

```
        }
              }
              else
              {
      if (VerifyFacts(conceptObj, conceptObjSon) == FALSE)
      { AddSon(conceptObj, conceptObjSon)
      if (currLevel <= finalLevel)
      GetOffspring(conceptObjSon, relationObj, ++currLevel, finalLevel)
              }
          }
      }
}
```

It is a recursive method of *ConceptTreeClass* that attaches to the tree object the tree built with the children.

```
ConceptTreeArray::GetDescendingTree(TermClass termObj, RelationClass
relationObj, integer finalLevel)
{
        integer level = 0
        ConceptClass conceptRoot
        if (conceptRoot.GetConceptfromTerm(termObj)!=0)
              END
        for all contexts of relationObj
        {
                ConceptTreeClass conceptTree;
                conceptTree.AddNode(conceptRoot)
                conceptTree.GetOffspring(conceptRoot, relationObj, level,
                finalLevel)
        AddTree(conceptTree)
        }
        return conceptTreeArray;
}
```

It reconstructs all the offspring concept trees (one for each context of the involved relation) for a given *ConceptClass* object derived by a *TermClass* object and a *RelationClass* object.

REFERENCES

1. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inform Med 1998;37:394–403.
2. Cimino JJ. From data to knowledge through concept-oriented terminologies: experience with the medical entity dictionary. J Am Med Inform Assoc 2000;7:288–97.
3. Ingenerf J, Giere W. Concept-oriented standardization and statistics-oriented classification: continuing the classification versus nomenclature controversy. Methods Inform Med 1998;37:527–39.
4. Musen MA, KE Wieckert, Miller ET, Campbell KE, Fagan LM. Development of a controlled medical terminology: knowledge acquisition and knowledge representation. Methods Inform Med 1995;34:85–95.
5. Lindberg D, Humphrey B, McCray A. The unified medical language system. Methods Inform Med 1993;32:281–9.
6. Bernauer J, Schoop D. Formal classification of medical concept description: graph-oriented operators. Methods Inform. Med. 1998;37:510–7.
7. Niemann H, Sagerer GF, Schroder S, Kummert F. Ernest: A semantic network system for pattern understanding. IEEE Trans Pattern Anal Machine Intell 1990;12(9):883–905.
8. Côté RA, Rothwell DJ, Patolay JL, Beckett RS, Brochu L, editors. SNOMED international: the systematized nomenclature of human and veterinary medicine. Northfield, IL: College of American Pathologist, 1993.
9. Schultz EB, Price C, Brown PJB. Symbolic anatomic knowledge representation in read codes version 3: structure and application. J Am Med Inform Assoc 1997;4:38–48.
10. Rector A.. Compositional models of medical concepts: towards

re-usable application independent medical terminologies. In: Barahona P, Christensen JP, editors. Knowledge and decisions in health telematics. Lisbon: IOS Press, 1994; 109–14.

11. Rector AL, Gangemi A, Galeazzi E, Glowinski AJ, Rossi Mori A. The GALEN CORE model schemata for anatomy: towards a re-usable application-independent model of medical concepts. In: Barahona P, Veloso M, Bryant J, editors. Proceedings of the 12th International Congress of the European Federation for Medical Informatics (MIE 94), Lisbon: IOS Press, 1994, 229–33.
12. Rector AL, Rogers JE, Pole P. The GALEN high level ontology medical informatics Europe '96, Brender J. *et al.*. editors. Lisbon: IOS Press, 1996, 174–8.
13. Rector AL, Bechhofer S, Goble CA, Horrocks I, Nowlan WA, Solomon WD. The GRAIL concept modeling language for medical terminology. Artif Intell Med 1997;9:139–71.
14. Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinshaw KP, Brinkley JF. Motivation and organizational principles for anatomical knowledge representation. J Am Med Inform Assoc 1998; 5:17–40.
15. Schubert R, Pommert A, Höhne KH. Methods for model-based knowledge representation in anatomy. In: Annual Symposium on Computer Applications in Medical Care, Proc SCAMC, New Orleans, LA, 1995, 971–7.
16. Ackerman, MJ, Spitzer, VM, Scherzinger, AL, Whitlock, DG. The visible human data set: an image resource for anatomical visualization. Medinfo 1995;8:1195–8.
17. Spence AP and Mason EB. Human anatomy and physiology. New York: West Publishing Co., 1992.
18. Balboni GC, *et al.* Anatomia umana. Milan, Italy:Edi-ermes, 1990.
19. Schiemann T, Nuthmann J, Tiede U, Höhne KH. Segmentation of the Visible Human for high quality volume based visualization, In Höhne KH, Kikinis R, editors. Visualization in biomedical computing, Proc VBC 96, Lecture Notes in Computer Science 1131, Berlin: Springer-Verlag, 1996, 13–22.
20. Campbell KE, Das AK, Musen MA. A logical foundation for representation of clinical data. J Am Med Inform Assoc 1994;1(3):218–32.
21. Brinkley JF, Bradley SW, Sundsten JW, Rosse C. The digital anatomist information system and its use in the generation and delivery of web-based anatomy atlases. Comput Biomed Res 1997;30:472–503.
22. Pommert A, Schubert R, Riemer M, Schiemann T, Tiede U, Höhne KH. Symbolic modeling of human anatomy for visualization and simulation. In: Robb RA, editor. Visualization in biomedical computing. Rochester, MN: Proc SPIE 2359, 1994, 412–23.